THG Graphics Card Buyers Guide

The THG Graphics Card Buyer's Guide has been written to be a guideline for the purchase of a new graphics card. It aids beginners in selecting the right model with the right feature set, and explains the newest technologies and features.

#1: Intended Use

#2: Technology

#3: Performance & Image Quality

#4: Budget

#5: Manufacturer & Feature Set

#6: The Purchase

Buying a new graphics card may seem like a simple matter at first. After all, both Internet shops and local retail stores carry a plethora of graphics cards in every performance and price category. This large variety of cards, however, makes it hard to select the one that is ideal for you. A multitude of factors need to be considered in the selection process, to ensure that the choice you make will keep you happy over as long a period of time as possible.

This article covers all of the criteria involved in selecting and buying the graphics card that is right for you. How important each factor is will depend on your personal preferences and the way you intend to use the card. For example, some people will require a video-in line, and for them this will be a make-or-break feature; others will not care about this particular capability. To help you define your requirements, we will also give a short overview of the technologies used in graphics cards of the past and present.

We've broken this buyer's guide up into six large sections that cover all of the important factors. Obviously, there is no perfect way to prioritize selection criteria, because preferences and needs differ for each individual. The order that we present here is only one possibility among many, and is meant more as a guideline to help you find your own personal ranking of criteria. Remember also that it's sometimes difficult to draw a line between these issues, so there will be some overlap in certain areas.

#1: Intended Use
A Short Overview

No matter what the intended use of your PC, be it games, office work, photo and video editing or anything else, you're going to need a graphics card. However, the importance of the card's performance depends greatly on the nature of the application! These days, the most important differentiating factors are video and 3D performance and quality.

The first step in determining your ideal graphics card is to take stock of the primary applications for which you use your PC. If most of your time on the computer is spent using office applications (word processing, spreadsheets), or other 2D software, then the 3D performance of a graphics card won't play a great role in your buying decision.

However, in future operating systems such as Microsoft's "Longhorn", the user interface will make much heavier use of a graphics card's 3D functionality, so 3D performance may be potentially important even for those who do not use 3D applications. For example, to use even the simplest 3D version of the Longhorn interface -- which goes by the name "Aero" -- full DirectX 9 support and 32MB of video memory are likely to be the bare minimum graphics card requirements. The grander "Aero Glass" interface version will require DirectX 9 support and 64MB of video memory!

Of course, there is still some time until Longhorn makes it to the marketplace and a computer near you. And even when it arrives, it will also come with a 2D-only user interface for systems that don't meet the 3D requirements. You can get more info on Microsoft's Longhorn here:

http://www.microsoft.com/whdc/device/display/graphics-reqs.mspx.

There are measurable 2D performance differences between individual cards and the various chip generations. However, the 2D performance of current graphics processors has reached such a high level overall that these differences won't make a tangible difference in everyday use, for example in a Windows XP environment. Applications such as Word, PowerPoint, Photoshop or Acrobat won't run any faster on a bleeding-edge high-end card than on a mainstream offering. This means that these days, a graphics card's performance is determined nearly entirely by its 3D performance.

Modern games such as Doom3 are very demanding on graphics cards.

Since today's graphics cards differ the most in 3D performance, this is the probably the main factor to look for if you intend to do any gaming on your PC. The variety of different card models from different generations and price brackets is enormous, as are the differences in 3D performance and feature sets. Even if you're more of a casual gamer that only plays a game every now and then, you shouldn't try

to save money in the wrong place. After all, gaming time is your free time, and you don't want to ruin it with stuttering or low-detail graphics. Cut too many corners and may end up with more exasperation than entertainment.

The 3D architecture of the card -- that is, which generations of which 3D standards it supports -- is very important. Usually, adherence to 3D standards is expressed in terms of support for a certain generation of Microsoft's DirectX 3D API, which is updated reg-

ularly. We'll talk about this some more later on in this guide. For now, we'd just like to mention that while most DirectX 8 compliant cards will be sufficient for current games, they won't do as well in the most recent and soon-to-come hit games, such as Doom III, Stalker and Half-Life 2.

If you're looking to replace your motherboard as well as your graphics cards, integrated graphics solutions may be an option for you. Beware, however, that the 3D performance of these solutions is, at best, comparable to that of the slowest add-in cards. As a result, these motherboards are only of limited use to PC gamers. If your focus lies more in the areas of office work and video editing, then they will usually be quite sufficient.

Recently, many companies have begun campaigns to secure a foothold for the PC in the living room. The primary selling point of such a solution is the PC's inherent suitability for video and audio playback. Again, special attention is given to the graphics card here as well. In principle, any graphics card is capable of displaying any video format, but there are major differences between cards in the resulting CPU load on the PC, and the output image quality. If the CPU load is too high when playing high-resolution HDTV videos (for example), there will be noticeable stuttering during playback. Graphics processors also differ in their offered color fidelity, and features such as de-interlacing and scaling. We'll look at this in more detail in section #2.

#2: Technology (Future Proofing)

Over the past few years graphics processors have evolved from pure 3D accelerators that could only perform pre-determined, specialized tasks, into real processors that are programmable to a certain extent. This development has allowed game designers to create their own 3D effects, in the same way as the creators of professional 3D rendering applications. These applications use their own programs for 3D effects, called *shaders*.

Simply put, a shader is a specified mathematical definition or description of an effect. For example, if a stone in a game is supposed to look wet, then a shader can be written for this purpose, which would define the sheen effect, reflections, incidence of light, and so on. The graphics processor then uses the shader to calculate this effect in real time.

In the past, the solution might have been taking the texture of the stone and overlaying it with a second texture that incorporates pseudo reflections, thereby creating the illusion of shininess. Of course, this wouldn't exactly have looked realistic. Today, these effects can be rendered with a high level of realism. In short, shaders add a great deal of realism to any game, though due to the topic's complexity, we will only be able to cover the most important aspects of how they work.

As we discussed earlier, a very important factor to consider when choosing a graphics card is which DirectX generation the graphics processor supports. The DirectX support of a card has important implications for its ability to make use of shaders, because each generation of DirectX increases the complexity of what calculations can be done by shaders. So, let's get back to the matter of DirectX generations.

DirectX Technology

DirectX 7 Class

The 3D engine of the game Battlefield 1942 sits solidly on a DirectX 7 foundation. Through the clever use of textures, the developers really squeeze

a lot out of the engine, but the in-game world is very static; dynamic lighting is not possible, for example. Another very popular DX7 game is Counter-Strike.

Games such as Quake 3 (OpenGL), Unreal, and even comparatively recent games such as Battlefield 1942 belong to this generation. Almost all effects in these games are realized through simple textures. Aside from transformation and lighting (T&L), these cards are not programmable. In fact, not all graphics processors of this generation even offer T&L support; for example Intel's integrated i865G or ST Micro's Kyro II.

DirectX 8 Class

Unreal Tournament 2003 uses a number of DirectX 8 shader effects. As a result, the game's graphics look much better than that of older games, and the in-game world seems more alive.

Graphics processors truly began to become programmable starting with DirectX 8. There are two capabilities that need to be taken into account here, namely pixel and vertex (=geometry) calculations through shaders. DirectX 8 incorporated several different pixel shader models (SMs), which support varying levels of programmability (PS 1.0, 1.1 and 1.2 are part of DirectX 8, while PS 1.4 was added in DirectX 8.1). At first, the complexity of the shader programs was quite limited, but their complexity has increased with the newer shader models. There is only one vertex shader model that is shared by both DirectX 8 and DirectX 8.1: Vertex Shader 1.0.

Direct X 9 Class

FarCry can be considered the first game that makes consistent use of shaders. Thanks to DirectX 9, the surfaces look very realistic and react to changes in lighting, throw believable shadows, and more. The game's environment seems very "alive."

Microsoft's current 3D API is DirectX 9, which permits even more freedom in shader programming than DirectX 8, and also allows for longer and more complex shaders. It also introduces the floating-point data model, which allows for detail calculations that are much more exact.

ATI and NVIDIA are the two companies that dominate the consumer 3D market, and their cards offer varying levels of precision. While ATI's processors use 24-bit precision across the board, NVIDIA's cards also support 16-bit and 32-bit floating point modes (as well as some other FF formats). The rule of thumb here is simple: "the higher the precision, the more complex the calculation."

Which data format is required depends greatly on the effect that is to be created -- not every effect requires the highest available precision.

DirectX 9 also incorporates several pixel shader models. First there is the original SM 2.0, to which the evolutionary SM 2.0a and 2.0b were later

Introduction to DirectX 8:
http://msdn.microsoft.com/library/en-us/dndrive/html/directx112000.asp?frame=true

Programmable Shaders for DirectX 8:
http://msdn.microsoft.com/library/en-us/dndrive/html/directx01152001.asp?frame=true

Introduction to DirectX 9:
http://msdn.microsoft.com/msdnmag/issues/03/07/DirectX90/toc.asp?frame=true

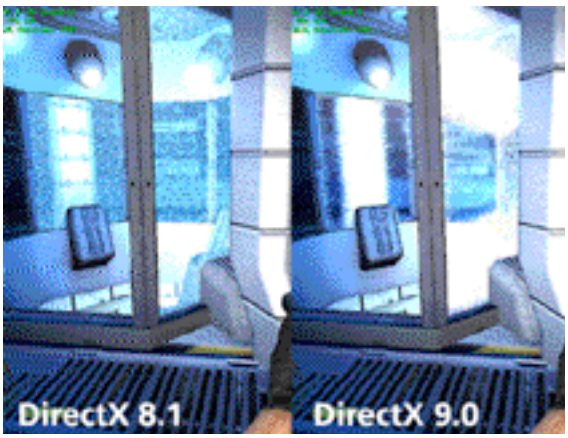
Shader Model 3.0:
http://www.microsoft.com/whdc/winhec/partners/shadermodel30_NVIDIA.mspx

Microsoft DirectX Overview:
http://msdn.microsoft.com/library/

added. SM 3.0 is a completely new and very recent addition, which is supported starting with DirectX 9.0c. Currently, only NVIDIA's GeForce 6xxx line of graphics processors can make use of SM 3.0.

If you would like to find out more about the various DirectX versions and the associated shader models, you will find lots of relevant information at the following sites:

It is important to note that you can't fully assess the graphics of a game solely by the DirectX version it uses. For example, DirectX 8 shaders can be used to implement many of the effects used these days,



which can bring even cutting-edge graphics processors to their knees. Game developers strive to use as low a DirectX version as possible, so they can target as large an audience as possible. How much computing power a shader will end up needing depends primarily on its complexity. Finally, it should also be noted that all cards are downward compatible. Upward compatibility is only possible in the case of vertex shaders which can be calculated by the CPU, and while possible, this would be very slow.

Two screenshots of the same scene in the game FarCry; one on a GeForce 4 Ti (DX8.1) and one on a GeForce 6800 (DX9).

Bear in mind that although many entry-level cards are DirectX 9 compliant, they are unable to deliver playable frame rates due to their low processing power (more on this in section #3). In some cases, the DirectX 9 compliance also refers only to certain areas. A prime example of this is Intel's new i915G integrated graphics chipset. Although the graphics processor supports Pixel Shader 2.0 (making it DirectX 9 compliant), it offloads all vertex shader calculations to the CPU, increasing CPU load.

OpenGL

After DirectX, OpenGL is the next most popular 3D API. It has existed for far longer than DirectX, and is available for a large number of operating systems. DirectX, on the other hand, is confined to Microsoft platforms.

Like DirectX, OpenGL is constantly being refined, updated and extended in its capabilities. Also like DirectX, it is supported by virtually every current 3D graphics card. Furthermore, the newest 3D features can usually also be implemented in OpenGL, even if these features have not yet been defined in the OpenGL standard; these are called *OpenGL extensions*. Frequently, graphics chip makers will offer their own extensions in drivers for certain effects that can be employed by applications or games. The two industry heavyweights, ATI and NVIDIA, offer very good OpenGL support, so there's not much to worry about there. Things aren't quite as rosy in the case of XGI and S3, however, which still have some room for improvement in their drivers.

Despite the seeming dominance of DirectX titles, there are still many games that are programmed for OpenGL. The most well known among these are the titles published by the Texan game designer id Software; many other game developers have also licensed 3D game engines from id to use in their own software. The newest and definitely most demanding OpenGL game from id is the first person shooter Doom III. NVIDIA cards perform especially well running this game, closely followed by ATI's offerings. The game will also run on XGI cards, with some effort and at reduced quality settings. For its part, S3 has published a special Doom III driver.

Interested readers can find more information on OpenGL at **http://www.opengl.org/**

Other Operating Systems

Things get more complicated for operating systems other than Microsoft Windows. The various cards' 3D performance under Linux differs drastically from that in Windows. Both

More information on Linux and graphics cards:

ATI Linux Drivers FAQ
(http://www.ati.com/products/catalyst/linux.html)

HOWTO: Installation Instructions for the ATI Proprietary Linux Driver
(http://www.ati.com/support/infobase/linuxhowto-ati.html)

NVIDIA Linux Advantage PDF
(http://www.nvidia.com/object/LO_20030328_6790.html)

NVIDIA Linux Driver Forum @ NVNews
(http://www.nvnews.net/vbulletin/forumdisplay.php?s=&forumid=14)

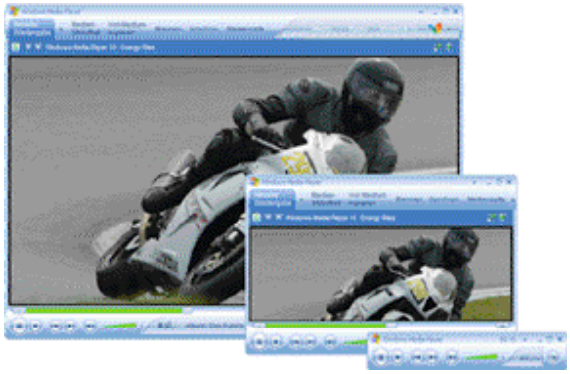
ATI and NVIDIA support Linux with special drivers.

Linux drivers can be found on ATI's and NVIDIA's download pages.

Video Playback

Video playback and Media Player visualizations can be accelerated by graphics cards, taking load off the CPU.

As we mentioned near the beginning of the article, video can be played back on practically any graphics card, as long as the correct codec is installed. Almost all graphics cards available today also offer special video acceleration features that handle effects such as resizing a video to fit a window, filtering and the like. The more tasks the graphics processor can handle, the less work is left to the CPU, improving overall performance. In the case of HDTV videos using very high resolutions, it is possible that the CPU alone isn't up to the task of decoding and playing back a video at all -- and this is where the video processor can step in to help.



Video acceleration is also an important issue for notebooks, as a CPU usually requires more power than a graphics processor. As a result, a good video acceleration will do its part in lengthening the running time of a notebook. Video acceleration features also come into play when watching DVDs.

Recently, both ATI and NVIDIA have put special emphasis on video features, and practically every new generation of graphics processors comes with extended video functionality. ATI groups together these capabilities, which can be found in the new X800 and X700 line of cards, under the name "FullStream HD." More information is available here: **http://www.ati.com/products/brochures/5639fullstreamWP.pdf.**

NVIDIA has equipped its newest chip family, the NV4x line, with a special, programmable video processor. This ensures support even for future video formats. Additionally, the video processor is designed to take some of the burden off the CPU when recording videos or during video encoding processes. More detailed information is available here: http://www.nvidia.com/object/feature_on-chip-video.html.

#3 Performance & Image Quality

Performance

The performance of a graphics card is normally measured by its frame rate, which is expressed in frames per second (FPS). The higher the frame rate a card can support, the more fluid the gaming experience will seem to the user. Essentially, a game displays a sequence of individual images (frames) in rapid succession. If they are output at a rate exceeding 25 fps, then the human eye is usually no longer capable of distinguishing the individual frames. However, in fast-paced games, such as first person shooters, even 25 fps will not be enough to make the game and all movements seem fluid. The bar for such games should be set at least at 60 fps.

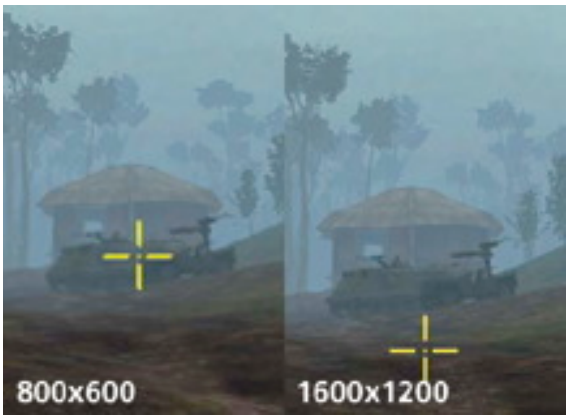
Aside from features such as FSAA and AF (which we will come to shortly), frame rate primarily depends on the selected screen resolution. The higher the resolution, the more pixels are available to display the scene, making the resulting output much more detailed. However, with increasing resolution, the amount of data that a graphics card has to handle also increases, meaning greater demands are placed on the hardware.

There are two important factors in assessing the ability of a graphics processor to provide high frame rate. The first is its pixel fill rate, which determines how many pixels can be processed per second (megapixels per second). The second is memory bandwidth, which measures how quickly the processor can read and write data from memory. In both cases, the "more is better" mantra applies.

At higher resolutions, more pixels are available to depict a more detailed image, as you can see in this image. While only very rough details can be made out at 800x600 (the small tree next to the Jeep), the detail level is much higher at 1600x1200.

Today, 1024x768 pixels is considered the standard

gaming resolution. The most popular higher resolutions are 1280x1024 and 1600x1200. In the case of classical CRT (cathode ray tube) monitors, the resolution can be selected freely, as long as it doesn't exceed the maximum possible physical resolution supported by the screen. Things are more complicated when TFT (thin film transistor, aka flat screen or LCD) monitors are used, since these have fixed resolutions. Any setting that differs from the monitor's native resolution requires that the image be interpolated, meaning either shrunk or enlarged.



Depending on the model that is used, this can have a noticeably adverse effect on image quality. Therefore, it is a good idea to choose a graphics card that offers good frame rates at your TFT's native resolution.

In addition to the resolution chosen, a card's frame rate will also depend to a great extent on the game being run. The extensive use of complex shaders in new games slows down many older cards unacceptably, even if these same cards offer very reasonable performance when running older titles. Most PC games allow for a reduction in detail level, thereby also reducing the number and complexity of effects, but this of course has a negative impact on the image quality and, consequently, on the gaming experience. The most important factor here is the DirectX support of both graphics card and game, which should be on the same level (see the section on DirectX Technology).

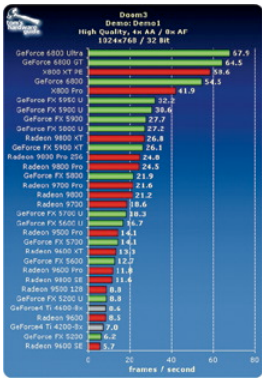
Benchmark Results

Since the performance of a card depends to such a great extent on the game being played and the selected resolution, a large number of combinations must be tested to reach a conclusive verdict on a card's performance. Cards from different manufacturers may show different performance in the same game.

This picture shows a typical benchmark table from the THG VGA Charts. Here, the game Doom3 was tested at a resolution of 1024x768 at 32-bit color depth. 4xFSAA and 8x anisotropic filtering were enabled, and the quality setting "High" was selected.

To determine a card's in-game performance, frame rate measurements are taken at distinctive points in the game. Many titles offer a recording feature for motion sequences, making it very easy to take comparable measurements for a number of cards. Some games measure the frame rate using a built-in function, while others require additional add-on utilities such as FRAPS. Another option for benchmarking tests is using in-game cut scenes, which are of course identical every time. Finally, for games that don't offer any of the choices above, the only remaining option is to try to replicate the same series of movements manually on every card.

The results found in the benchmark tables are usually the average of several tests, showing the average frame rate a card is able to sustain in a game. Thus, a result of 60 fps means that the frame rate may dip below and rise above that number at different places in the game. Minimum scores would be more meaningful, but these are very difficult to determine; dips in frame rate can be caused by in-game loading or background activity of the operating system, and these factors cannot be easily replicated. Therefore, the average frame rate remains the most meaningful measuring standard.



Despite this, we can't emphasize often enough that you need to remember that these are indeed *average* values. If a card only runs a game at an average of 25 fps, the game will show pronounced stuttering during its "slower periods" which may seem to turn it into a slide show. In general, you should be on the safe side with a card that pushes 60-100 fps in games – at the highest quality settings, of course.

You can find a good overview of the perform-

ance of different current and previous-generation graphics cards in the Tom's Hardware VGA Charts:

CPU

The system CPU has quite a bit of influence on the graphics card's performance. Even though modern graphics processors no longer need any CPU time for their calculations, the data they process has to be prepared by the CPU and then transferred to the card. Additionally, the CPU also must take care of handling computer player AI, physics calculations and sound, all at the same time. To be able to push a fast graphics card to its limit, you'll also need a potent CPU.

Of course, the opposite case is just as true – a fast processor won't do any good if the graphics card is limiting the frame rate. And the same also holds true for the system memory, which can hold the system back if it's too slow, or if there isn't enough of it. In summary, the individual components need to be well-balanced. A single weak component can cripple the entire system.

Fortunately, there aren't any bad choices where the graphics interface is concerned. The current standard is the AGP 8x bus, which will gradually be supplanted by its successor, PCI Express, over the coming months and years. For now, don't expect to see any performance increases from switching to the new bus, however! If you'd like to read up on PCI Express and its future role in the graphics market, take a look at our article here: **http://graphics.tomshardware.com/graphic/20040310/index.html.**

Comparisons with older graphics cards can be found in previous iterations of our VGA Charts:

http://www.tomshardware.com/graphic/20020418/index.html

http://www.tomshardware.com/graphic/20030120/index.html

http://www.tomshardware.com/graphic/20021218/index.html

http://www.tomshardware.com/graphic/20031229/index.html

http://graphics.tomshardware.com/graphic/20041004/index.html

FSAA and AF

The abbreviations FSAA and AF stand for two methods of improving the image quality in 3D games. FSAA is short for Full Scene Anti Aliasing, which is a technique for smoothing the edges of 3D objects within a scene. AF is shorthand for Anisotropic Filtering, which is a filtering method applied to textures on 3D objects to make them look crisper and less washed-out, greatly enhancing image quality. Both FSAA and AF are very demanding on graphics processors, especially when used in combination.

These features can usually be enabled or disabled through the graphics driver's control panel. Some games also let you enable them directly through the in-game options menu, without the need for special software. However, some games have trouble with FSAA, due to peculiarities of the graphics engine they use. In these cases, leaving FSAA disabled is usually the better choice, as image corruption can occur otherwise.

The advantages of FSAA become especially obvious on slightly slanted vertical object borders.

Anisotropic filtering results in much crisper textures.

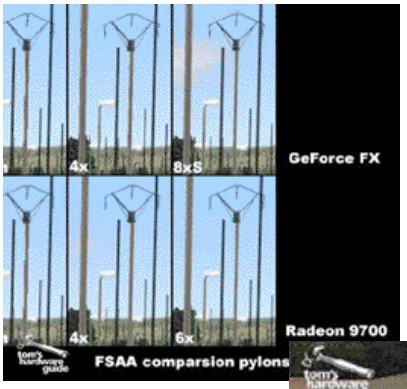
Although the underlying principles are the same everywhere, the technical implementation of these techniques differs from company to company and even from one card generation to the next. On older graphics cards or newer low-end models, FSAA can only be used to a limited extent; this is usually either because the card's performance is too low to deal with the extra calculations, or because it uses a slow or outdated FSAA method. There are also a number of different AF methods that differ both in calculation complexity and resulting image quality.

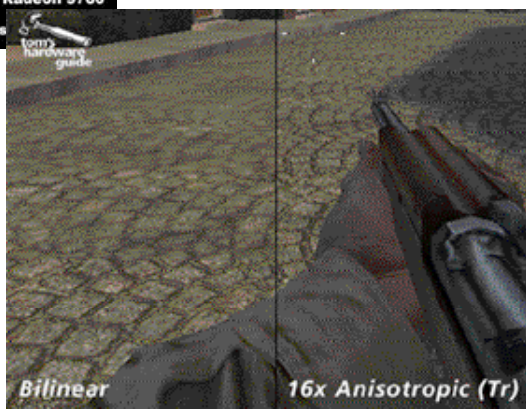
Both FSAA and AF require a lot of computing power and memory bandwidth. For this reason, ATI and NVIDIA use heavily "optimized" versions of these methods to achieve better results (higher performance) while still offering greatly improved image quality compared to the standard rendering output. The heaviest optimization is done on the anisotropic filtering implementations. As a result, there are some cases in which a reduction in image quality compared to the "correct" or "real" method becomes visible. Unfortunately, both of the big players like to use this method of tweaking too much in order to try to win benchmark comparisons. Therefore, image quality and performance can differ immensely between driver versions even on the same card!

You can read up on the texture filtering "optimizations" currently in use in the following article: **http://graphics.tomshardware.com/graphic/20040603/index.html**

Image Quality

Image quality is a topic that would easily merit its own article, if not a book in its own right. What I mean here is the quality of the rendered 3D scene as it appears on the player's screen. This whole discussion was originally caused by the tricks and tweaks that graphics card makers have begun to build into their drivers. Their goal is to get the most performance out of their cards, and to this end, sometimes certain calculations are either skipped or simplified. In principle, this is possible in a lot of places without the player being forced to accept reduced image quality. Unfortunately, the chipmakers tend to do a bit too much tweaking, especially to win performance comparisons. The result is often visibly reduced image quality, noticeable at least to experienced users. Casual gamers, on the other hand, may often not even notice anything. In our article (http://graphics.tomshardware.com/graphic/20040603/index.html) we took a look at a number of optimizations used by the graphics chip companies, and explained how they work and what effect they have on image quality and 3D performance.





Here is an image quality comparison taken from the game FarCry using older drivers. In this driver, NVIDIA replaced some of the game's own shaders with highly optimized ones. The result is visibly reduced image quality on the one hand, but improved performance on the other.

Meanwhile, the chipmakers have learned that many users don't necessarily want such optimizations, especially if they are forced upon them. Anyone who pays $500 (or more) for a graphics card understandably expects the highest possible image quality. This is especially so considering that such optimizations are not really that essential -- the enthusiast cards are now more than fast enough to handle the highest quality settings. In response, NVIDIA and ATI now allow for most of these optimizations to be switched off in their most recent drivers.

Another reason for reduced image quality can be the use of reduced floating-point precision in DirectX 9 games. A good example of this is the game FarCry. NVIDIA's GeForce FX cards render most of the shaders using only 16-bit precision, which leads to pronounced visual artifacts (see also: **http://graphics.tomshardware.com/graphic/20040414/geforce_6800-46.html).** While NVIDIA has addressed these quality issues with newer drivers, the result is that the frame rates have taken a nosedive as a result **(http://graphics.tomshardware.com/graphic/20041004/vga_charts-08.html).** NVIDIA was only able to overcome this performance handicap in DirectX 9 games with the new GeForce 6xxx line.

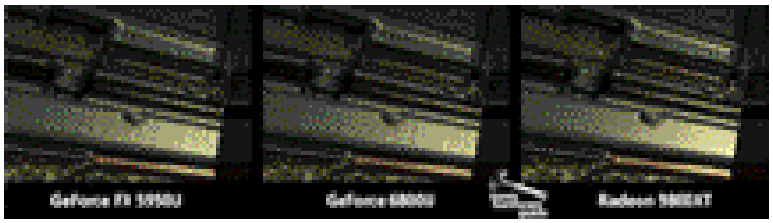
Since the image quality produced by a card can change with literally every driver release, we recommend staying informed by reading the reviews of new card generations, as we also regularly test the image quality in these articles.

#4 Budget (Card Overview)

Each graphics chip maker develops products for every price category. Pictured here is NVIDIA's roadmap from the year 2003.

Cards can generally be categorized into three large groups, each of which can once again be subdivided into two subgroups. The two big graphics chip companies, ATI and NVIDIA, offer different chips for each of the various price brackets. Note that the boundaries between the categories tend to blur quite a bit, however, due to price fluctuations in the market.

The three main price groups are the entry-level



or budget line, the mid-priced or mainstream products, and finally, the higher-end enthusiast cards. Again, within each of these there are two versions offering different performance levels -- one is the standard version, while the other runs at higher clock speeds. ATI denotes these faster cards by the addition of a "Pro" or "XT" to the card name, while NVIDIA's nomenclature uses the "GT" and "Ultra" suffixes.

Low-cost products are often tagged as SE or LE parts. However, these budget cards sometimes don't carry any special tag at all, making them hard to tell apart from "the real deal". In these cases, only careful attention to the technical data will help you from mistakenly purchasing the wrong card.

For some further reading about image quality, check out these articles:

http://graphics.tomshardware.com/graphic/20040603/index.html

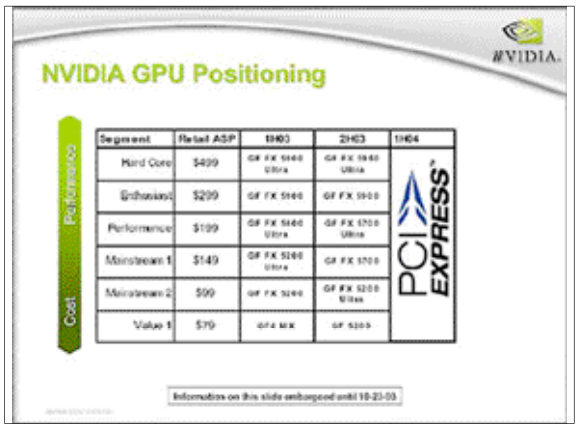
http://graphics.tomshardware.com/graphic/20040414/geforce_6800-43.html

http://graphics.tomshardware.com/graphic/20040504/ati-x800-32.html

NVIDIA is a chipmaker only, focusing its attention solely on designing and producing graphics processors, while leaving the production and sale of retail cards to its board partners. ATI, on the other hand, is quite active in the retail market as well, albeit only in the United States and Canada. Its cards are usually designated "Built by ATI", while those produced and sold by other companies are "Powered by ATI."

Another factor further complicating any attempt to categorize the cards by price alone are the graphics cards from older generations, which keep getting cheaper due to the introduction of newer models. There are especially pronounced differences between NVIDIA and ATI here. ATI's second to last generation of chips (Radeon 9500, 9700, 9800) is still very much up-to-date from a technological perspective, with DirectX 9 support and multisampling FSAA. Only the Radeon 9000 and 9200 cards are the exception here, as they are still based on the DirectX 8 design of the Radeon 8500 along with its slower super sampling FSAA implementation. Shader Model 3.0 is not supported by any ATI card at this point. The only cards that actually can take advantage of it are those of NVIDIA's GeForce 6xxx line.

In contrast, NVIDIA's second to last generation of cards are, by today's standards, technologically outdated (DirectX 8 and multi sampling FSAA on the GeForce 4 Ti, DirectX 7 on the GeForce 4 MX). The last iteration of the GeForce FX 5xxx series performed very well in DirectX 8 titles, but drops to mediocre levels in current DirectX 9 games. As mentioned before, this weakness has been corrected in the new GeForce 6xxx line (note the absence of the "FX" designation).



Price Categories

Let's now take a look at the three main price categories. We begin with the cheapest cards, which are the entry-level or low-budget products. These fall either into the sub-$100 category, or the price bracket between $100 and $150. The second category, usually called the "mainstream", begins at $150 and reaches up to the $300 mark. In this category, the largest selection of cards can be found between $150 and $250. Last, we have the enthusiast category which starts at around $300 and extends to $500 (and well beyond, in some cases.) This is where the latest top models from ATI and NVIDIA are to be found.

In the following overview, we have also listed cards from older generations that are still available in the market. The prices quoted here are current as of mid-October 2004; we take no guarantee for the correctness of this information.

Note that in some cases it is rather difficult to determine which models actually exist in the mar-

ket and what specifications they use. The low-cost sector, especially, is flooded with a multitude of different configurations for the same basic chip. A good starting place to get an overview is Gigabyte's product page **(http://tw.gigabyte.com/VGA/Products/Products_ComparisonSheet_List.htm).**

Older Radeon Models

Radeon 9200

The RV 280 (Radeon 9200), like its predecessor the RV 250 (Radeon 9000), is based on the DirectX 8.1 design of the Radeon 8500 (R 200). Compared to the Radeon 8500 with its 4x2 pipe design, this chip only features half as many texture units per pixel pipeline (4x1) and only one vertex shader unit. The main differences between the Radeon 9000 and the 9200 are the newer part's higher clock speeds, and its support for the AGP 8x interface. It is produced on a 0.15µ process and contains roughly 32 million transistors.

The greatest weaknesses of the Radeon 9200 are its outdated and slow super sampling FSAA implementation, as well as it being limited to bilinear filtering.

Versions:

Radeon 9200 SE - 64/128 MB - 64-/128-bit DDR - 200/330 MHz
Radeon 9200 - 64/128 MB - 64-/128-bit DDR - 250/400 MHz
Radeon 9200 PRO - 128 MB - 128 Bit DDR - 300/600 MHz

Radeon 9600

The Radeon 9600, which has the internal designation RV350, is the successor to the highly successful DirectX 9 chip RV300 (Radeon 9500). The RV300 only differed from the "big" R300 (Radeon 9700) in that it featured a memory bus that was pared down from 256 bits to 128 bits. In the standard version of the chip, ATI also disabled four of the eight pixel pipelines. Nonetheless, it was the exact same chip as the R300; its approximately 107 million transistors made it expensive to produce as a mainstream part. In the newer RV350, ATI didn't just disable some of the pixel pipes through the card's BIOS, but physically reduced the number to four in the chip design. Combined with a die-shrink to a 0.13µ process, this made the 75-million transistor chip much cheaper to produce.

The Radeon 9600's advantage over its predecessor lies in its much higher clock speeds, which usually outweighs the disadvantages incurred by the reduction in the number of pixel pipelines. Despite this, the Radeon 9600 Pro is sometimes outperformed by the Radeon 9500 Pro in fill-rate intensive applications. Other than that, the 9600 offers DirectX 9, modern multi-sampling and fast anisotropic filtering – in short, everything that the flagship products have.

The Radeon 9600XT (codename RV360) takes a special place in this line-up, though, as it is based on a more modern architecture than the earlier 9600 variants. For the first time, this makes driver optimizations for trilinear filtering possible, which results in much higher performance.

Versions:

Radeon 9600 XT - 128/256 MB - 128bit - 500/600 MHz
Radeon 9600 Pro - 128/256 MB - 128 Bit - 400/600 MHz
Radeon 9600 - 64/128/256 MB - 128 Bit - 325/400 MHz
Radeon 9600SE - 64/128 MB - 64/128-bit - 325/365 MHz

Articles:

http://graphics.tomshardware.com/graphic/20030416/index.html
http://graphics.tomshardware.com/graphic/20031015/index.html

Radeon 9800

ATI's flagship model of the past few years carries the internal designation R350. The main change from its predecessor, the Radeon 9700 (code name R300), is the increased clock speed, resulting in improved performance (especially when FSAA and AF are enabled). While other details were changed and improved as well, these aren't really noticeable in practice. The chip is produced on an 0.15µ process and consists of 107 million transistors. Its advantage over its smaller siblings lies in its 256-bit memory interface, giving it a higher memory bandwidth, and a full complement of eight pixel pipelines. During the product run, ATI also introduced a 256MB version featuring DDR II video memory.

Entry-Level

Price Range	Bus	Lowest Price	Model	Memory	Memory Bus	Shader Model	Pixel Pipes
<$99	AGP	$35	XGI Volari V3	64MB	64-bit	1.3	2
		$45	XGI Volari V3	128MB	64-bit	1.3	2
		$46	NVIDIA GeForce FX 5200	128MB	64-bit	2	4
		$55	NVIDIA GeForce FX 5200	128MB	128-bit	2	4
		$56	ATI Radeon 9550 SE	128MB	64-bit	2	4
		$60	ATI Radeon 9600 SE/LE	128MB	64-bit	2	4
		$63	Matrox Millennium G550	32MB	64-bit	DX6	2
		$64	ATI Radeon 9550	128MB	128-bit	2	4
		$65	ATI Radeon 9600 SE/LE	128MB	128-bit	2	4
		$70	NVIDIA GeForce FX 5500	128MB	128-bit	2	4
		$79	ATI Radeon 9550	256MB	128-bit	2	4
		$80	NVIDIA GeForce FX 5500	128MB	64-bit	2	4
		$80	ATI Radeon 9600	128MB	128-bit	2	4
		$80	NVIDIA GeForce FX 5200	256MB	128-bit	2	4
		$92	NVIDIA GeForce FX 5500	256MB	128-bit	2	4
		$95	NVIDIA GeForce FX 5700 LE	256MB	128-bit	2	4
	PCIe	$75	ATI Radeon X300 SE	128MB	64-bit	2	4
		$77	NVIDIA GeForce PCX 5300	128MB	64-bit	DX7	2
$100-$149	AGP	$100	ATI Radeon 9600	256MB	128-bit	2	4
		$106	NVIDIA GeForce FX 5700	128MB	128-bit	2	4
		$110	ATI Radeon 9600 Pro	128MB	128-bit	2	4
		$130	NVIDIA GeForce FX 5200U	128MB	128-bit	2	4
		$125	ATI Radeon 9600 Pro	256MB	128-bit	2	4
		$131	ATI Radeon 9800 SE	128MB	128-bit	2	4
		$140	ATI Radeon 9600 XT	128MB	128-bit	2	4
	PCIe	$105	ATI Radeon X300	128MB	128-bit	2	4
		$110	NVIDIA GeForce PCX 5750	128MB	128-bit	2	4
		$135	NVIDIA GeForce 6600	128MB	128-bit	3	8

Main Stream

Price Range	Bus	Lowest Price	Model	Memory	Memory Bus	Shader Model	Pixel Pipes
$150-$199	AGP	$160	Matrox Millennium P650	64MB	128-bit	1.3	2
		$161	ATI Radeon 9600 XT	256MB	128-bit	2	4
		$164	NVIDIA GeForce FX 5700 Ultra	128MB	128-bit	2	4
		$175	NVIDIA GeForce FX 5900 SE/XT	128MB	256-bit	2	8
		$195	Matrox Millennium P750	64MB	128-bit	1.3	2
	PCI	$150	ATI Radeon X600 Pro	128MB	128-bit	2	4
		$170	NVIDIA GeForce 6600	256MB	128-bit	3	8
		$175	ATI Radeon X600 XT	128MB	128-bit	2	4
		$180	ATI Radeon X700 Pro	128MB	128-bit	2	8
$200-$299	AGP	$200	ATI Radeon 9800 Pro	128MB	256-bit	2	8
		$215	NVIDIA GeForce FX 5900	128MB	256-bit	2	8
		$250	ATI Radeon 9800 Pro	256MB	256-bit	2	8
		$270	NVIDIA GeForce 6800	128MB	256-bit	3	12
		$288	Matrox Parhelia 128	128MB	256-bit	1.3	4

Enthusiast

Price Range	Bus	Lowest Price	Model	Memory	Memory Bus	Shader Model	Pixel Pipes
$300-$399	AGP	$303	ATI Radeon 9800XT	256MB	256-bit	2	8
		$350	NVIDIA GeForce 6800 GT	256MB	256-bit	3	16
		$369	NVIDIA GeForce FX 5900 Ultra	256MB	256-bit	2	8
		$370	NVIDIA GeForce FX 5950 Ultra	256MB	256-bit	2	8
		$380	ATI Radeon X800 Pro	256MB	256-bit	2	12
	PCIe						
$400-$499	AGP	$465	ATI Radeon X800 XT	256MB	256-bit	2	16
	PCIe						
>$500	AGP	$525	NVIDIA GeForce 6800 Ultra	256MB	256-bit	3	16
		$550	Matrox Parhelia 256	256MB	256-bit	1.3	4
		$680	ATI Radeon X800 XT PE	256MB	256-bit	2	16
	PCIe	$560	ATI Radeon X800 XT	256MB	256-bit	2	16

Not Available Yet (in USA)

Price Range	Bus		Model	Memory	Memory Bus	Shader Model	Pixel Pipes
	AGP		NVIDIA GeForce 6600	128MB	128-bit	3	8
			NVIDIA GeForce 6600	256MB	128-bit	3	8
			NVIDIA GeForce 6600 GT	128MB	128-bit	3	8
			NVIDIA GeForce 6800 LE	128MB	128-bit	3	8
			S3 S8 Deltachrome				
			XGI Volari V5 Series				
			XGI Volari V8 Series				
	PCIe		NVIDIA GeForce 6800U	256MB	256-bit	3	16
			NVIDIA GeForce 6800GT	256MB	256-bit	3	16
			NVIDIA GeForce 6800	128MB	256-bit	3	12
			NVIDIA GeForce 6600 GT	128MB	128-bit	3	8
			NVIDIA GeForce 6200	128MB	128-bit	3	4
			ATI Radeon X700 Pro	256MB	128-bit	2	8
			ATI Radeon X700 XT	128MB	128-bit	2	8
			ATI Radeon X800 Pro	256MB	256-bit	2	12
			ATI Radeon X800 XT PE	256MB	256-bit	2	16

clock speed, the smaller X800 Pro features only 12 pixel pipes. Basically, the architecture is an evolution of the Radeon 9600 XT, but with a great number of improvements added as well as extra features. These cards are available as AGP products, and, in some cases, PCI Express parts as well.

Versions:

Radeon X800 Pro - 12PP - 256MB - 256-bit - 475/900 MHz
Radeon X800 XT - 16PP - 256MB - 256-bit - 500/1000 MHz
Radeon X800 XT PE - 16PP - 256MB - 256-bit - 520/1120 MHz

Article:

http://graphics.tomshardware.com/graphic/20040504/index.html

Older NVIDIA Models

GeForce FX 5200

With the chip internally codenamed NV34, NVIDIA brought DirectX 9 to the low-cost market segment, replacing the outdated GeForce 4 MX line (DirectX 7). Like its bigger siblings, it features complete DirectX 9 support. However, NVIDIA reduced the number of pixel pipelines to four and didn't give the chip the modern memory interface of the bigger models. Instead, it uses the time-tested solution from the GeForce 4 Ti generation. The vertex shader performance is also reduced relative to higher-end models. The chip has a transistor count of about 45 million and is produced on a 0.15µ process.

In light of the very limited performance and the only moderate clock speeds, DirectX 9 support seems to be more of a paper feature than a real boon here. In practice, the chip is simply too slow for complex DirectX 9 calculations in resolutions of 1024x768 and above. Despite this, the chip is still quite a good performer for an entry-level card. This is due to the memory interface, the multi sampling FSAA, and the average (trilinear) filtering performance, inherited from the GeForce 4 TI cards. Beware of non-Ultra parts, though, as some of them are only equipped with much slower 64 bit memory modules.

Versions:

GeForce FX 5200 - 64/128/256 MB 64-/128-bit - 250/400 MHz
GeForce FX 5200 Ultra - 128 MB - 128-bit - 325/650 MHz

Articles:

http://graphics.tomshardware.com/graphic/20030311/index.html
http://graphics.tomshardware.com/graphic/200303061/index.html

GeForce FX 5600

This chip carries the internal designation NV31 and is produced on a 0.13µ process. It was meant to be the successor to the highly successful GeForce 4 Ti 4200 line. Shortly after its introduction near the beginning of 2003, NVIDIA improved the Ultra version of the card; thanks to the switch to a flip-chip design, NVIDIA was able to increase the clock speed by another 50MHz to 400MHz. The previous Ultras were then supposed to be sold as standard chips, but whether or not this was always the case is hard to tell. By now all of these remnants should be off the shelves, but there is no guarantee that the old chips might not still be found on Ultra cards. Prospective buyers should therefore keep an eye on the clock speeds. If your card only runs at a clock speed of 350MHz, it still carries the older version of the chip.

From a technological perspective, this DirectX 9 card features all the functionality of its bigger brother, such as Color Compression, fast (adaptive) anisotropic filtering and multi-sampling FSAA. Only the number of pixel pipelines fell victim to the "red pencil," leaving just four. Also, the card features a 128-bit memory interface instead of the high-frequency 128-bit DDR II memory of the NV30 (FX 5800) or the 256-bit memory of the NV35 (FX 5900).

Versions:

GeForceFX 5600 - 128 MB/256MB - 128 Bit - 325/550 MHz
GeForceFX 5600 Ultra - 128 MB/256 MB - 128 Bit - 400/700 MHz

Articles:

http://graphics.tomshardware.com/graphic/20030311/index.html
http://graphics.tomshardware.com/graphic/200303061/index.html

With the R360, aka Radeon 9800 XT, ATI once more extended the 9800 product line at the high end. Compared to the Radeon 9800 Pro, the XT ran at even higher clock speeds, and ATI also optimized the architecture. Radeon 9800 XT cards are only available with 256MB of video memory.

Beware of the Radeon 9800 SE, however. Unlike the rest of the 9800 family, this chip only features four active pixel pipelines and is therefore closer to a Radeon 9600. On top of that, the SE also features a trimmed-down 128-bit memory interface.

Versions:

Radeon 9800 SE - 4PP - 128MB - 128-bit - 380/675 MHz
Radeon 9800 - 8PP - 128 MB - 256 Bit - 325/580 MHz
Radeon 9800 Pro - 8 PP - 128 MB - 256 Bit - 380/680 MHz
Radeon 9800 Pro - 8PP - 256 MB - 256 Bit DDR II - 380/700 MHz

Articles:

http://graphics.tomshardware.com/graphic/20030306/index.html
http://graphics.tomshardware.com/graphic/20030604/index.html
http://graphics.tomshardware.com/graphic/20030930/index.html

ATI's Current Product Family

Radeon 9250

The Radeon 9350 is based on the Radeon 9200 series but operates at much lower clock speeds.

Versions:

Radeon 9250 - 128/256 MB - 128-bit - 240/250 MHz

Radeon 9550

From a technological viewpoint, the Radeon 9550 is nearly identical to the Radeon 9600.

Versions:

Radeon 9550 SE- 128MB - 64-bit - ??/?? MHz
Radeon 9550 - 128MB/256MB - 64-/128-bit - 250/400 MHz

Radeon X300

The Radeon X300 is the PCI Express version of the Radeon 9550.

Versions:

Radeon X300 SE - 64/128/256 MB - 64-bit - 325/400 MHz
Radeon X300 - 128/256 MB - 128-bit - 325/400 MHz

Radeon X600

The Radeon X600 line traces its technological roots back to the Radeon 9600XT. This card is only available as a PCI Express version.

Versions:

Radeon X600 Pro - 128MB - 128-bit - 400/600 MHz
Radeon X600 XT - 128MB - 128-bit - 500/740 MHz
Radeon X600 XT - 256MB - 128-bit - 500/600 MHz

Radeon X700

The Radeon X700 series carries the internal part name RV410 and replaces the seemingly short-lived X600 line. Technologically, the chip is based on the X800 (R420) design. As is the standard procedure for the mainstream chips, ATI has halved the number of pixel pipelines to eight and limited the memory interface to 128 bits. The number of vertex shader units remains unchanged at six.

Versions:

Radeon X700 - 128MB GDDR3 - 128-bit - 400/700 MHz
Radeon X700 Pro - 128MB/256MB GDDR3 - 128-bit - 425/860 MHz
Radeon X700 XT - 128MB GDDR3 - 128-bit - 475/1050 MHz

Article:

http://graphics.tomshardware.com/graphic/20040921/index.html

Radeon X800

The Radeon X800 cards, codenamed R420, constitute ATI's current high-end offering. While the X800 XT Platinum Edition (PE) and the X800 XT both feature 16 pixel pipelines and differ only in

GeForce FX 5900

Only a few months after the introduction of the previous top model, the 0.13µ GeForce FX 5800 (NV30), NVIDIA replaced the heavily criticized card (loud cooling solution, great heat dissipation, too low memory bandwidth) with the FX 5900 (NV35). In addition to re-designing the reference cooling solution so it is much quieter, NVIDIA also decided to drop the very hot DDR II memory on this card, instead widening the memory bus to 256 bits. 3D features saw only minor improvements or tweaks (Color Compression and floating-point performance, UltraShadow feature). Of note is the fact that the FX 5900 Ultra chip is clocked 50MHz slower than the FX 5800 Ultra. In exchange, the memory bandwidth grew from 16.7 GB/s to a very impressive 27.2 GB/s. The number of transistors also increased slightly, from about 125 million to 130 million.

Obviously, the FX 5900 Ultra is the fastest card of the family. Since the Ultra version is only available in a 256 MB configuration, it is also the most expensive of the bunch. 128 MB and a lower price might have made more sense in making the card more attractive. The non-Ultra version runs at slightly lower clock speeds, while the GeForce FX 5900 XT seems to offer the best price to performance ratio. Although running at lower frequencies than the two faster models, it offers the full feature set.

Versions:

GeForceFX 5900 XT - 128 MB - 256-bit - 400/700 MHz
GeForceFX 5900 - 128 MB - 256 Bit - 400/850 MHz
GeForceFX 5900 Ultra - 256 MB - 256-bit - 450/850 MHz

Articles:

http://graphics.tomshardware.com/graphic/20030512/index.html

NVIDIA's Current Product Family

GeForce FX 5500

The GeForce FX 5500 is based on the FX 5200. Aside from the change in clock speeds there are no known differences.

Versions:

GeForce FX 5500 - 128/256 MB - 64-/128-bit - 270/400 MHz

GeForce FX 5700

The GeForce FX 5700 series is based on the GeForce FX 5950 (NV38), but is cut down to four pixel pipelines. It features the same number of vertex shader units as its bigger brother, though. During the product cycle, NVIDIA refined the Ultra version of the FX 5700, giving it GDDR 3 memory.

Versions:

GeForce 5700 LE - 64/128MB - 64-/128-bit - 400/500 MHz
GeForce 5700 - 128/256MB - 64-/128-bit - 425/550 MHz
GeForce 5700 Ultra - 128MB - 128-bit - 475/900 MHz
GeForce 5700 Ultra - 128MB GDDR3 - 128-bit - 475/950 MHz

Articles:

http://graphics.tomshardware.com/graphic/20040405/index.html

GeForce FX 5950

The NV38, or GeForce FX 5950, is a further refinement of the NV35 architecture. The main improvement in this case is the higher clock speeds, as the card was meant to be an answer to ATI's Radeon 9800 XT.

Versions:

GeForce FX 5950 Ultra - 256MB - 256-bit - 475/950 MHz

Articles:

http://graphics.tomshardware.com/graphic/20031023/index.html

GeForce PCX 5300/5750/5900

The GeForce PCX series is NVIDIA's first product line for the PCI Express interface. The cards are based on the existing AGP versions, and the model numbers compare as follows: PCX 5300 = GeForce 4 MX, PCX 5750 = FX 5700, PCX 5900 = FX 5900. Note that the PCI Express versions run at different clock speeds from the AGP versions, however!

Versions:

GeForce PCX 5300 - 128MB - 64-bit - 250/332 MHz
GeForce PCX 5700 - 128MB - 128-bit - 425/500 MHz
GeForce PCX 5900 - 128MB - 256-bit - 350/550 MHz

GeForce 6200

Meant to be an affordable entry-level card, the GeForce 6200 rounds out the NV4x line at the bottom. At the very recent introduction of this PCI Express line of cards, NVIDIA used modified GeForce 6600 processors with some features disabled. It stands to reason that NVIDIA will use a different, newly designed chip to save on costs once the parts begin to ship to retail. Currently, an AGP version is not planned.

GeForce 6600

The GeForce 6600 (aka NV43) is the first mainstream line of cards built on the NV4x architecture. To reduce the production cost of the chip. NVIDIA reduced the number of pixel pipelines to eight, pared down the vertex shader units from 6 to 3, and slimmed down the memory interface to 128 bits. Two models have been announced so far: the GeForce 6600 GT and the 6600. NV43 is also NVIDIA's first native PCI Express part. According to NVIDIA, an AGP version of the 6600 using the HSI Bridge chip is already being prepared.

Versions:

GeForce 6600 - 8PP - 128/256MB - 128-bit - 300/550 MHz
GeForce 6600 GT - 8PP - 128/256MB - 128-bit - 500/1000 MHz

Articles:

http://graphics.tomshardware.com/graphic/20040812/index.html
http://graphics.tomshardware.com/graphic/20040907/index.html

GeForce 6800

The GeForce 6800 is the first product family of NVIDIA's NV4x line of chips. The fact that NVIDIA has dropped the "FX" from the name emphasizes how much the company is trying to distance itself from previous generations with this chip. The entire architecture has been thoroughly overhauled, and the weaknesses of the NV3x addressed. As a result, the NV4x cards no longer suffer a performance penalty when running DirectX 9 shaders at full floating-point precision. Also, with support for DirectX 9.0c and Shader Model 3.0, NVIDIA is one step ahead of the competition. Compared to the previous flagship models of the FX 59xx line, this card offers more than twice the performance. Unfortunately, the blazingly fast Ultra versions have very strict power supply requirements (more on this shortly).

The cards of the GeForce 6800 line are available in three versions. The Ultra version features 16 pixel pipelines, runs at clock speeds of 400/1100MHz (core/memory) and requires two auxiliary power connectors. The GT version differs from the Ultra only in that it runs at a slower speed: 350/1000MHz, and makes do with only one extra power connector. Finally, there is also the "vanilla" GeForce 6800 (without any suffix) which features only 12 pixel pipelines and 128MB of memory.

The GeForce 6800 GT and Ultra models are available in both AGP and PCI Express flavors, while the vanilla 6800 currently ships solely as an AGP part.

Versions:

GeForce 6800 - 12PP - 128/256MB - 256-bit - 325/700 MHz
GeForce 6800 GT - 16PP - 256MB - 256-bit - 350/1000 MHz*
GeForce 6800 Ultra - 16PP - 256MB - 256-bit - 400/1100 MHz*

*The PCI Express models support NVIDIA's SLI technology.

Articles:

http://graphics.tomshardware.com/graphic/20040414/index.html

Beyond NVIDIA and ATI - Alternatives

There are very few real alternatives to cards based on chips from ATI or NVIDIA. While boards using XGI or S3 chips are available, these don't have any real presence in the market. Only SiS has been comparatively successful in the market, thanks to the low-cost Xabre 400 and 600 cards. These cards have been plagued by driver problems and low texture quality, though, which should be considered before a purchase.

#5 Manufacturer & Feature Set

Once you have found a model that suits you, the time has come to choose the right cardmaker. As we mentioned earlier, NVIDIA, S3 and XGI don't sell cards themselves, choosing instead to focus their attention exclusively on the design and production of their graphics processors. While ATI sells cards, their scope of operation is limited to Canada and the USA ("Built by ATI"). ATI-based cards produced by other companies usually say "Powered by ATI".

Performance differences between cards using the same chip are the exception, rather than the norm. Cardmakers usually adhere quite strictly to the clock speed specifications suggested by NVIDIA, with a couple of exceptions. First, a few companies offer special "OC" versions in the enthusiast segment which run at higher clock speeds, and even most "normal" cards can be set to operate at higher clock speeds manually. The difference is that in the first case, the manufacturer guarantees that the card will work at the higher speed without a problem.

Graphics cards in the lowest-price segment are the second exception. Here, the competition is so intense that every dollar counts, so manufacturers may try to cut corners. Often enough, the result is either low 3D performance or bad 2D quality, or even both. Frequently, the memory bus is reduced in size, for example from 128 bits to 64 bits (see above). We can only warn you to stay away from such products, since the 3D performance suffers enormously when memory bandwidth is halved. If detailed information is conspicuously absent from the box or the salesperson can't or won't give you any technical data for the card, the best course of action is to keep looking, no matter how tempting the offer may be.

Another variation on the game of "sales poker" is the use of lures that sound good but may not necessarily be of great value. For example, while 256MB of video memory obviously sounds better than 128MB, the extra memory will only pay off in modern games with large textures, played at high resolutions and with FSAA and AF enabled. To be able to handle the amount of data produced in such a scenario, both the graphics processor and the memory bus need to be sufficiently fast. In other words, 256MB simply doesn't make any sense outside of the enthusiast segment!

Image quality is a factor that is very hard to judge before a purchase. To prevent nasty surprises, you should always ask the staff about the store's return policy before you buy the card. Then test it to see if it meets your needs, and return it if necessary.

AGP or PCI Express

Without a doubt, the future belongs to the new PCI Express interface. However, the technology is still too new to judge when exactly this future will be upon us; in other words, when PCI Express will become a "must-have". So far, only motherboards using Intel chipsets offer PCI Express at all, although the competition supporting AMD is set to go.

A typical AGP 8x slot (top) and the new x16 PEG (PCI Express for Graphics) slot.

The different connectors. AGP on top, PCI Express below.

Where 3D performance is concerned, PCI Express offers at best minimal advantages over AGP models. Buyers looking to upgrade their graphics card won't be making a mistake if they pick up an AGP model, assuming that their system is reasonably powerful and up-to-date. However, if the potential upgrade would also include a new motherboard and a CPU, it's worth taking a look at PCI Express. Keep an eye on the price tag, though, as PCI Express systems on the whole aren't significantly faster than their AGP counterparts at the moment. In the end, it's up to the individual to decide how much an investment into the future is worth to them.

We've compiled more information on PCI Express in the following article: http://graphics.tomshardware.com/graphic/20040310/index.html.

SLI

When older gamers hear the term "SLI", their eyes tend to glaze over and they wax nostalgic. These three letters bring back fond memories of the glorious times when the now-defunct 3D chipmaker 3dfx was still around. The abbreviation SLI stood for a technique that allowed two 3D cards to work in tandem in one computer, splitting the work between them. This led to a performance boost factor of somewhere between 1.5 and 2.

The AGP bus put an end to this type of solution, but now, with the introduction of PCI Express, SLI

is experiencing something of a revival with the help of NVIDIA. The new interface allows for several x16 PEG (PCI Express for Graphics) slots on one board. The success of NVIDIA's SLI technology will depend mostly on the pricing and the availability of motherboards with the appropriate support. So far, SLI capability has been announced for PCI Express versions of the GeForce 6800 Ultra, 6800 GT and 6600 GT cards. You can read up on SLI here: http://graphics.tomshardware.com/graphic/20040628/index.html.

Power Requirements

Modern graphics processors are very complex; the newest flagship models contain more than 200 million transistors. Currently, the record for most transistors in a consumer graphics chip is held by NVIDIA's GeForce 6800 GT and Ultra chips, which weigh in at 220 million transistors. To give you a frame of reference, bear in mind that even Intel's Pentium 4 EE CPU consists of "only" 178 million transistors – 149 million of which make up the second level cache!

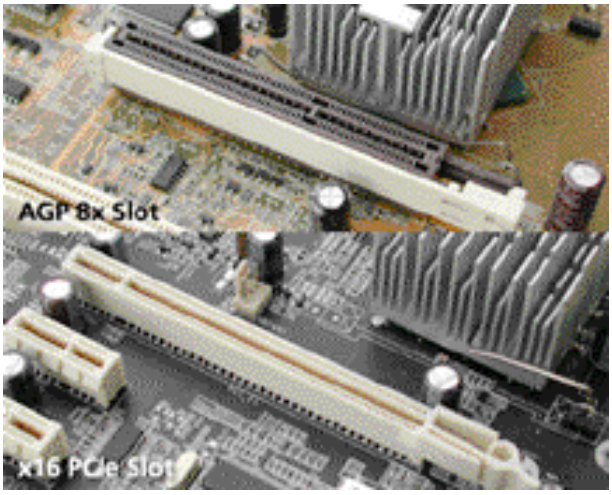
The hunger for power of modern graphics cards is correspondingly high, and can no longer be satisfied through the current supplied through the AGP slot alone. While the AGP slot supplies up to 45 watts, a GeForce 6800 Ultra draws up to 110 watts under full load. To make up for the difference, 2 additional ATX Molex connectors need to be plugged into the card. This is an extreme example, and most cards need only one such auxiliary power plug. The new PCI Express interface improves this situation by offering up to 75 watts, but even this obviously isn't enough for the highest-end cards.

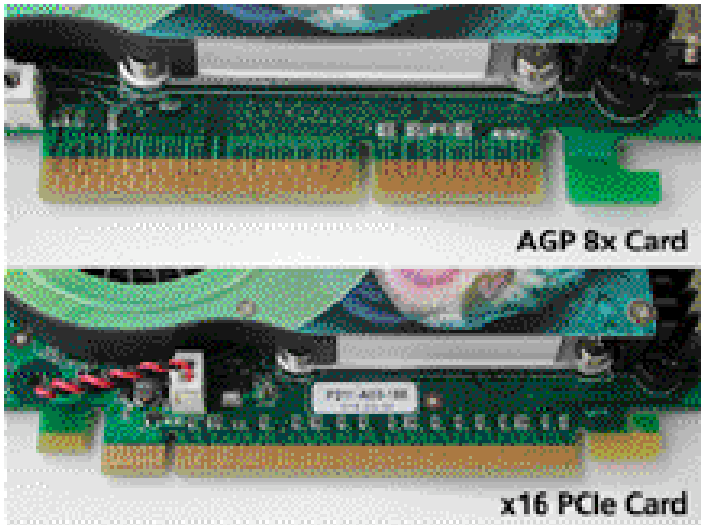
The way the auxiliary power cables are split up is important. Whenever possible, the graphics card should always have its own, unshared cable. The only other component that can be attached to the same cable without risking trouble is a case fan.

Depending on the power requirements of a card, the power supplied by the motherboard may not be enough. In this case, cards require auxiliary power connectors fed directly by the system's power supply. This picture shows the types of connectors currently in use.

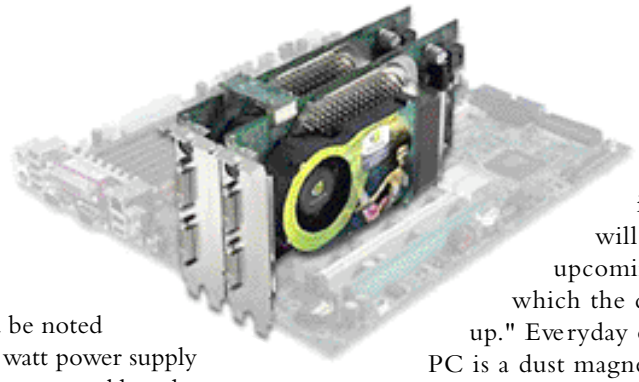
In 2D mode, graphics cards draw comparatively little power. However, in a 3D game that puts a lot of stress on the graphics card, the CPU and potentially even the hard drive, the power draw can peak quite suddenly and overwhelm the PSU. The unavoidable consequence is a crash of the entire system.

As a reaction to the power needs of their cards, both ATI and NVIDIA state minimum requirements for power supplies. However, these should only be considered guidelines, at best, since the power supply also needs to power the CPU, the drives and every other component in the system. The power requirements quoted by the manufacturers refer to a standard PC with a modern CPU, say a Pentium 4 3.2GHz, a hard drive, a DVD drive and a soundcard. If your computer houses more





components, it's better to play it safe, bite the bullet and buy a brawnier PSU.



Also, it should be noted that a good 350 watt power supply can deliver much more stable voltages than a cheap 450 watt model. The specifications of the PSU, such as the amperage at a certain voltage, can be of great help in making an educated buying decision. Such information can usually be found on the manufacturer's website. If the manufacturer doesn't publish any information, it's usually safer to choose a different model. You can find more information on power supplies in this article: http://www.tomshardware.com/howto/20040122/index.html.

Power requirements of the newest high-end models:

X800 XT PE: min. 350 watts and one additional connector on an unshared cable
X800 Pro, 9800 XT: min. 300 watts and one additional connector
GeForce 6800 Ultra: min. 350 watts and two additional connectors; cable can be shared. 480 watts required for overclocking, with two unshared cables.
GeForce 6800 GT und 6800: min. 300 watts. One unshared cable from the power supply.

Smaller models are less demanding where the power supply is concerned. A standard 300 Watt power supply will usually be sufficient, at least as long as there aren't too many other components in the system.

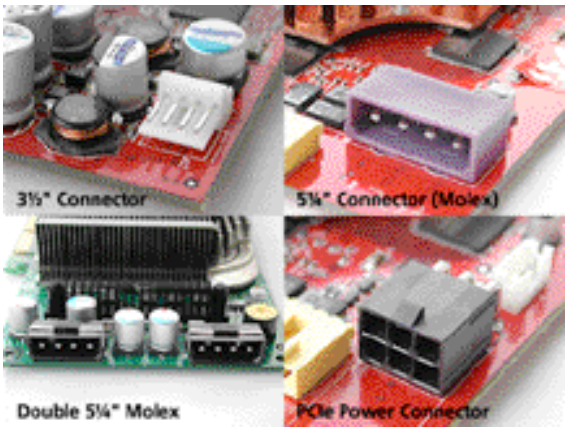
Looks & Cooling

Features that make a card stand out visually, such as colored PCBs, classy materials or colorful fan LEDs, should be considered gimmicks; after all, the cards are currently installed into the motherboard upside-down anyway. Often, such extras will only make the installation more difficult. If, for example, the card features a large heat sink on the back, the result can be a cramped motherboard or, in the worst case, conflicts with existing parts. Sometimes, the cards' length can also be a problem.

Everyday Life: The card's spiffy fan is no longer visible once it is installed into the system. This will only change with the upcoming BTX case standard, in which the cards are installed "right side up." Everyday occurrence: the inside of a PC is a dust magnet.

More important than its looks is a cooler's efficiency and noise level. Thankfully, most manufacturers have picked up on the trend towards quieter PCs and have begun equipping their cards with very quiet, temperature-controlled fans. The reference fan (i.e. the cooling design suggested by the chip designer) is often a good choice. You should only choose a graphics card with a different cooler if the cardmaker in question also offers detailed information on the differences in noise level and cooling efficiency compared to the standard cooling solution.

On entry-level cards that usually feature graphics chips running at low clock speeds, a passive cooling solution is often sufficient. The advantages: no noise and absolute reliability.

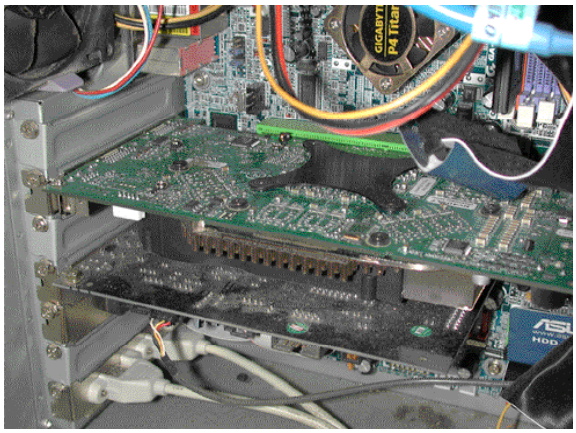


Many companies needlessly equip their entry-level cards with cheap and simple fans that are usually loud and short-lived. Often, a passive cooling solution would do just as well, and last longer.

In the case of slower graphics cards, it's a good move to chose a model that features passive cooling, since the fans found on cards in this price category are usually more loud than helpful. The high-end segment features passively cooled cards as well, though. However, the computer case needs to have good airflow for them to work well.

Monitor Connectivity

Almost every graphics card today features connectors for two monitors. Usually, cards will come with one DVI-I connector for a digital monitor and one standard VGA connector for CRTs. Adapters that ship with the card also allow for a second analog monitor to be used instead of a digital one. So, typically, the following combinations are possible: 1x VGA, 1x DVI, 2x VGA or 1x DVI & 1x VGA. If you're planning to hook up two digital (TFT) monitors to your card, you should look for cards with two DVI-I connectors; these are becoming more and more popular. Their flexibility allows for



practically any combination of monitors.

If you intend to use more than two monitors, you will either have to buy a more expensive workstation card, or take a look at Matrox's selection of multi-head cards. ATI offers another alternative in the form of its IGP 9100 motherboard chipset sporting integrated graphics. Thanks to its SurroundView feature, the on-board graphics can remain active even when an add-in card is installed in the motherboard. As a result, up to three displays can be attached to one computer. However, for gaming purposes, simply hooking up several monitors to your system won't be enough. You can find a little guide here: **http://graphics.tomshardware.com/graphic/20040216/index.html.**

Another factor that is an unknown at present is the impact of PCI Express, or more precisely, motherboards with several x16 PEG slots. Several such chipsets have already been announced but have not yet been released by their respective manufacturers.

At present, only a few graphics cards come with two DVI(-I) monitor outputs, for connection with digital flat-panel monitors. A combination of 1x VGA and 1x DVI(-I) connectors is more common. Dual DVI-I is the more future-proof choice, even if you only attach analog monitors at present (using adapters – see next picture).

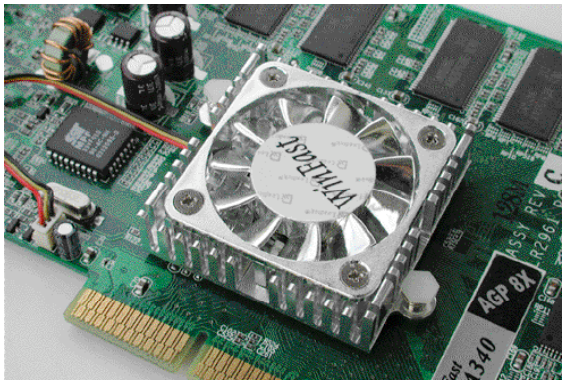
Using special adapters, analog monitors can be hooked up to DVI-I connectors. In most cases, a graphics card will ship with such an adapter.

Video In & Out

Nearly all cards come equipped with video-out functionality, but video-in features are much rarer. Often video-in is only present on so-called "ViVo" models, which offer connection options for video sources via composite (RCA video) or S-VHS cables. However, these are of no help for connecting digital video cameras, as these tend to require a FireWire jack that is only rarely found on a graphics card. You can only watch television via video-in if the source signal is coming from a tuner, i.e. a video recorder, satellite receiver or TV.

As an alternative, graphics cards with an integrated TV tuner are also available. ATI offers its All-in-Wonder series, while NVIDIA sells its Personal Cinema line. Of course, these cards cost more than the standard models. Also, bear in mind that you will lose all of the functionality come the next graphics card upgrade (unless you buy a newer version of these cards, again at an additional expense.) If you consider yourself a frequent upgrader, you should probably opt for an add-in TV card instead.

It is important to understand that even cards with video-in functionality are not "real" video capture, cutting and/or editing cards. Although modern cards offer hardware-supported encoding of video



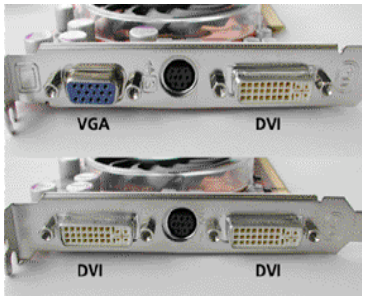
material, the brunt of the burden is still borne by the CPU!

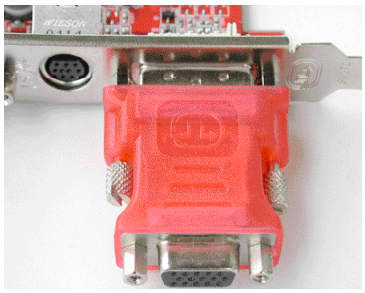
Most cards with video-in & out connect to video sources using breakout boxes or dongles. Cameras or video recorders can also be connected.

Video-out in HDTV quality is a new and hot topic. While most new graphics chips support this feature, the appropriate RGB cable is (still) usually absent from the bundle. If this is an important feature for you, be sure to check the cards' feature and accessory table.

Software Bundle

One significant area that sets individual card makers apart from each other is the software they bundle with their cards. Aside from the obligatory driver CD, most cards come with a software DVD player. Graphics cards with video inputs also often come with video editing software. In most cases, these programs are either older or slimmed-down versions, usually carrying the SE or LE tag.





Some companies also develop their own display tools that operate in parallel with the graphics driver. These aren't required, though, since the card makers have no part in driver development these days; they are written exclusively by the chip maker.

Depending on the card manufacturer, some games may even be included in the bundle. These run the gamut from entirely useless (outdated titles, limited versions or freely available demo versions) to highly attractive (retail versions of top titles). You can save money this way, if you find a card that ships with a game you were planning on buying anyway. In most cases, the bundled versions ship without a manual, however. And of course, there's no advantage if you've already bought the game, or if it isn't one you like.

As usual, it's up to the individual to decide how much of a factor the software and/or gaming bundle is in the buying decision.

#6 The Purchase

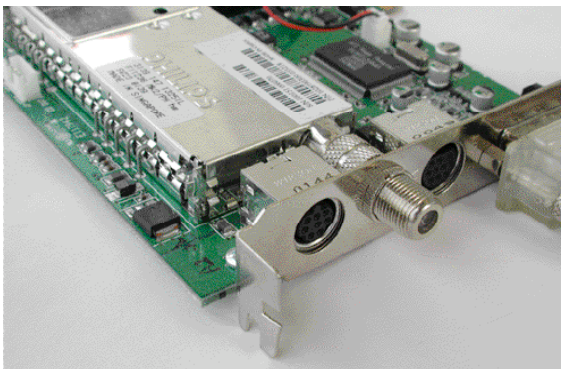
Once you've picked out the right model, it's time to clear the next hurdle, namely the purchase itself. Again, there is a lot to consider. First, you need to decide whether you want to order online or buy at a local store. On the whole, online shops tend to offer lower prices, but make sure to check out the cost of shipping and handling! Many online retailers charge a premium, and that would-be bargain can turn out to be more expensive than it would have been at your local store.

Buying at a store can also offer some other advantages, especially if you can get competent advice from the employees. If you're lucky, your local store may even let you test the card in the showroom, so you can check out the 2D signal quality of the VGA output, for example. This is an especially important factor in the case of low-cost models.

The main thing to remember is to closely scrutinize the particulars of the offer in question. If information on the card's memory and core frequencies is conspicuously absent, it's best to keep looking! An exact listing of the card's specifications is the least a customer should expect. Frequently, even a scan of the manufacturer's spec sheet may not be of any help. Especially in the low-cost sector, many manufacturers just don't give detailed information on their cards' clock speeds or memory bus width, as the following pictures show:

Abit's data sheet for the Radeon 9200SE-T shows quite detailed information on the card's specifications, with only the clock speeds missing. The note about the slower 64-bit memory is particularly important.

MSI does the exact opposite. Not only are the



clock speeds not to be found anywhere, a reference to the slow 64-bit memory is also absent. Instead, the spec sheet only gushes about the memory size of 128MB - which has practically no beneficial impact on the performance of a card in this category.

If you already have all of the information you need about your dream card, you can search for the best prices online by using search engines such as Bizrate (**http://tomshardware.bizrate.com/buy/browse__cat_id--4.html**). If you've already had a positive experience with an online shop, it may pay off to check if they have the card you're looking for, even if it is slightly more expensive there. Also, always make sure to check the availability of the card! Often, shops will list cards that they don't have in stock. In the worst case, you may end up waiting for your card for weeks or even months.

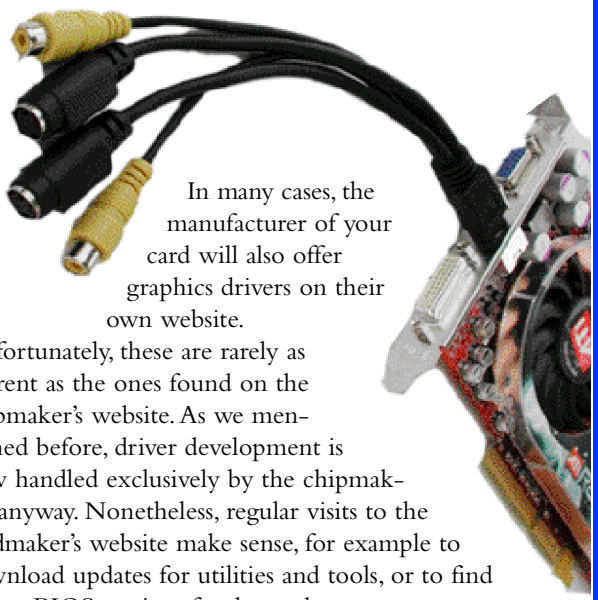
If you're unsure whether the card you've picked out is the right model, would like to take a look at the cardbefore you buy or have questions about the installation of the card, you're better off going to a store – assuming you find one that offers competent advice.

Drivers

The drivers on the CD that comes with your new card will probably already be out of date by the time you buy it. So, after you get the card home, we recommend downloading the latest drivers for it from the Internet. Newer games, especially, tend to have more problems with older drivers. The same holds true for Microsoft's DirectX, which you should also keep up to date. Newer games often ship with a runtime installer of the current DX version, though. To make life a little easier for you, we have compiled a list of the most important links:

Current Drivers:

ATI-based graphics cards (Radeon, All In Wonder, IGP) - http://www.ati.com/support/driver.html
Intel integrated graphics (chipsets i865G, 915G) - http://downloadfinder.intel.com/scripts-df/support_intel.asp?iid=HPAGE+header_support_download&#
NVIDIA-based graphics cards (GeForce, nForce) - http://www.nvidia.com/content/drivers/drivers.asp
S3-based graphics cards (Deltachrome) - http://www.s3graphics.com/drivers.jsp
SIS integrated graphics - http://download.sis.com/
XGI-based graphics (Volari) - http://www.xgitech.com/sd/sd_download.asp



In many cases, the manufacturer of your card will also offer graphics drivers on their own website. Unfortunately, these are rarely as current as the ones found on the chipmaker's website. As we mentioned before, driver development is now handled exclusively by the chipmakers anyway. Nonetheless, regular visits to the cardmaker's website make sense, for example to download updates for utilities and tools, or to find newer BIOS versions for the card .

Closing Words:

The Tom's Hardware Graphics Card Buyer's Guide is meant as a guide and a reference to help beginners select and buy a new graphics card. Through constant updates, it is our goal to create a comprehensive and current overview of the models available in the marketplace and the technologies upon which they are based. Be sure to drop by regularly – it will be worth your while! ■